An Investigation of the Trolley Problem in Regard to Automated Cars

Kurt Metz, Brandon Amdur, Dane Cipparone | Bryan Weaver, Justin D'Arms | The Ohio State University

Problem

- Automated vehicles are involved in fewer accidents than human-operated vehicles, but crashes are still inevitable.
- The trolley problem, shown above, involves a decision between saving 1 life or saving 5 lives. There are many variations of this problem that involve slightly different circumstances.
- With a human driver, quick decisions are made based on instinct and reaction, while automated cars base decisions on pre-determined algorithms.
- This raises concerns about what the cars should do in each scenario, and who should make that decision.

Significance

- Automated cars are already on the road.
- It's estimated that there will be as many as 10 million self-driving cars on the road by 2020.
- The National Safety Council estimates that **over 40,000 people died** in motor vehicle crashes in 2016.
    - First time since before 2008.

Methodology

For this project, the focus is mainly on the transitional phase from human-operated vehicles to 100% adoption of self-driving vehicles. Once automation is fully adopted, some of the concerns addressed in this project will no longer be issues.

In order to properly analyze and determine what an automated car **should** do, we first need to define what an automated car **can** do. For this project we made the following assumptions to represent the ideal automated car:

- Can recognize situations as well or better than humans
- Can analyze and react much faster than humans
- Can communicate with other automated vehicles
- SAE Level 5 (National Highway Traffic Safety Administration)

Possible Solutions

Utilitarian

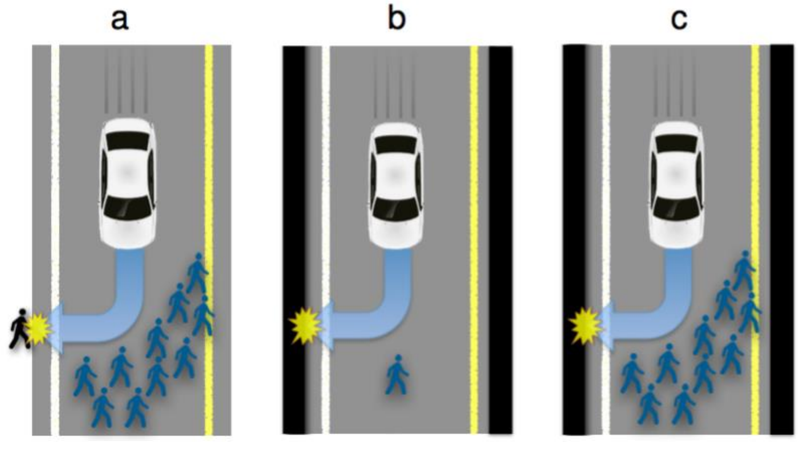
- Car prioritizes the decision that saves the most lives
- Pros:
    - Fewer casualties if possible for every crash
    - Ideally, this approach saves the most lives overall
- Cons:
    - People are less likely to buy a car that may choose to save someone else over them and their passengers
    - The increased safety of automated cars is not helpful if people aren't using them.

Self-preserving

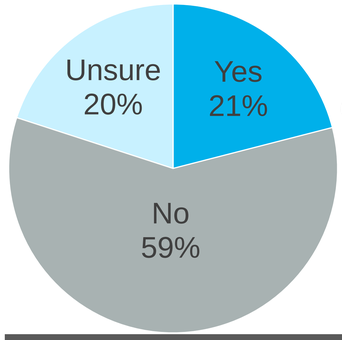
- Car prioritizes driver safety above all else
- Pros:
    - More appealing to consumers
    - More adoption of automated cars means safer roads due to less crashes in general
- Cons:
    - Possibility of fewer lives being saved in certain crashes
    - Morally objectionable/unethical

Random

- Car makes crash-scenario decisions randomly
- Pros:
    - Avoids responsibility of choosing who lives
    - Avoids possible bias
- Cons:
    - Death toll could be higher than other approaches
    - With the ability to save more lives, there is a moral obligation to do so



Would Buy Automated Car if Utilitarian



"People want to live in a world in which driverless cars minimize casualties, but they want their own car to protect them at all costs" – *Iyad Rahwan*

Our Stance

Our group heavily favored a utilitarian approach because saving the most lives is important to us. To prevent bias, the government would need to set crash-specific regulations that car manufacturers would be required to follow.

To be street legal, cars from all manufacturers must behave the same in a crash. Cars would also have to meet certain standards for sensor quality to ensure the roads are as safe as possible. Eventually, it would become illegal to use any car that is not automated, except in designated areas, such as rural areas.

We understand that prohibiting human driving is not a popular solution, but we believe that it is the necessary action to save lives. The self-preserving and random approaches both sacrifice safety for the sake of consumer popularity and comfort. Ultimately, the purpose of self driving cars themselves should be supported: saving the most lives.

This transition will likely be lengthy, and we may not see 100% adoption of automation in our lifetime. However, it is still important to have an open discussion about legislation regarding self-driving cars before they become a mainstream consumer product.

Acknowledgements

We want to thank our advisors, Bryan Weaver and Justin D'Arms, for giving us this opportunity and helping us prepare for this forum.

We also want to thank the OSU Philosophy Department, the College of Arts & Sciences, and the Battelle Foundation for providing us with the resources and funding necessary to complete this research project.